RESEARCH COMMUNICATION

Localization of the receptor-binding site in the collectin family of proteins

Rajneesh MALHOTRA,*‡ Steen B. LAURSEN,† Antony C. WILLIS* and Robert B. SIM*

*Medical Research Council Immunochemistry Unit, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, U.K. and †Institute of Medical Microbiology, Bartholins Building, University of Aarhus, University Park, 8000 Aarhus, Denmark

Collectin receptor (C1q receptor) has been shown to bind human C1q, mannose-binding protein (MBP), lung surfactant protein A (SP-A) and bovine conglutinin. These ligands have a similar ultrastructure, each consisting of collagenous and globular domains, but do not show a high degree of sequence similarity. For C1q and SP-A, it has been shown that both bind to cell-surface-expressed receptor(s) via their collagenous regions and this is likely to be the case with the other ligands. Within the collagenous region, near the 'bend' region of the collagen triple helix in C1q, MBP and SP-A, a cluster of similar charged residues is observed. This region has been suggested to be associated with receptor binding. A similar region of charge density occurs close to the N-terminus of conglutinin. In this paper we describe a truncated form of conglutinin, which has 55 amino acids missing from the N-terminus and does not bind to the collectin receptor. The results presented here strongly indicate that receptor–ligand interaction is mediated via the N-terminal region of conglutinin, consistent with the earlier proposal for the binding site.

INTRODUCTION

Collectins are a group of soluble proteins each of which has collagenous and non-collagenous domains, and may be complement associated and/or have lectin activity [1]. Collectins include lung surfactant protein A (SP-A), mannose-binding protein (MBP), C1q, lung surfactant protein D (SP-D) and conglutinin. All the molecules belonging to the collectin family consist of multiple polypeptide chains each made up of a short non-collagenous N-terminal segment, followed by a region of collagen-like sequence (characterized by the repeating triplet sequence Gly-Xaa-Yaa, where Yaa is often a hydroxylated amino acid). The C-terminal portion of each polypeptide is non-collagenous, and contains, in SP-A, MBP, conglutinin and SP-D, a structure known as a C-type lectin domain, which indicates that these proteins exhibit calcium ion-dependent binding to carbohydrates. In C1q, the C-terminal halves of the polypeptide sequences contain an immunoglobulin-binding domain. The subunits of the proteins are built up by association of three polypeptide chains, the collagenous regions of which intertwine to form a collagen triple helix. In SP-A, C1q, conglutinin and MBP, the collagen triple helix is bent in each subunit, due to an interruption in the repeating Gly-Xaa-Yaa sequences of the polypeptide chains. The non-collagenous C-terminal halves form a globular 'head'. Six such subunits, in SP-A, C1q or MBP, associate to form the characteristic 'bunch of tulips' structure seen in electron microscopy. In this structure, the globular heads form the 'flowers', and the collagen helices form the 'stalks' (Figure 1a).

At least four collectins (C1q, MBP, SP-A and conglutinin) bind to a single receptor (collectin receptor) [2]. Binding of SP-D to the collectin receptor has not been firmly established. The collectin receptor (also called C1q receptor) is a 56 kDa glycoprotein [3]. The N-terminal amino acid sequence of isolated collectin receptor and the sequence of peptides obtained by

V8/trypsin digestion show a high degree of similarity to the cDNA-derived amino acid sequence of a human protein reported as a component of RoSSA [4] or as calreticulin [5]. The comparison of immunological, composition, cellular localization and peptide-sequence data between purified collectin receptor and an isolated RoSSA component or the RoSSA cDNA sequence indicates that collectin receptor and RoSSA are similar, but not identical, molecules [6]. Collectin receptor has been isolated from a number of cell types, including human tonsil lymphocytes, the monocytic cell line U937 [3], endothelial cells and platelets [7] and the lung alveolar type-II cell line A549 [8]. Although the binding characteristics of collectin–collectin receptor interaction are well documented [2,9], the site of binding between collectin and collectin receptor has not been investigated. The ligands of the collectin receptor have similar quaternary structures, but are not highly homologous at the primary structure level. We therefore examined the known sequences of the ligands to select regions of similarity, which might be involved in the receptor binding. Experimental data on collectin–collectin receptor interaction was then obtained to confirm localization of the binding site.

MATERIALS AND METHODS

SDS/PAGE

SDS/PAGE was carried out as described by Laemmli [10]. Samples were prepared (reduced or alkylated) as described in [11]. Proteins were detected with Coomassie Blue staining.

Protein purification

C1q was isolated from human serum as described by Reid [12]. Collagen stalks from C1q were generated by pepsin digestion of C1q and purified by the method described by Reid [12]. Human collectin receptor was purified from human tonsil lymphocytes

(a)

Polypeptide chains

Subunit structure

Conglutinin (T)

Cleavage site

Conglutinin (N)

SP-A or C1q

(b)

		1								10										20										30
Conglutinin	(N)	A	E	M	T	T	F	S	Q	K	I	L	A	N	A	C	T	L	V	M	C	S	P	L	E	S	G	L	P	G	H	D

		32													46													60
Conglutinin	(N)	G	Q	D	G	R	E	C	P	H	G	E	K	G	D	P	G	S	P	G	P	A	G	R	A	G	R	P	G	W

Conglutinin (N): positions 32, 46, 60, and below 1 ... 5

																												1				5
Conglutinin	(T)	_____ G R P G W

		18													32													43
Human C1q (A chain)	G	R	R	G	R	P	G	L	K	G	E	Q	G	E	P	G	A	P	G I R T G	I	Q	G

		24													38													51
Human C1q (B chain)	G Q P G T P G I K G E K G L P G L A G D H G E F G E K G

		18													32													45
Human C1q (C chain)	G K D G Y D G L P G P K G E P G I P A I P G I R G P K G

		35													49													60
Human MBP	G K D G R D G T K G E K G E P G Q G L R G L Q G P P

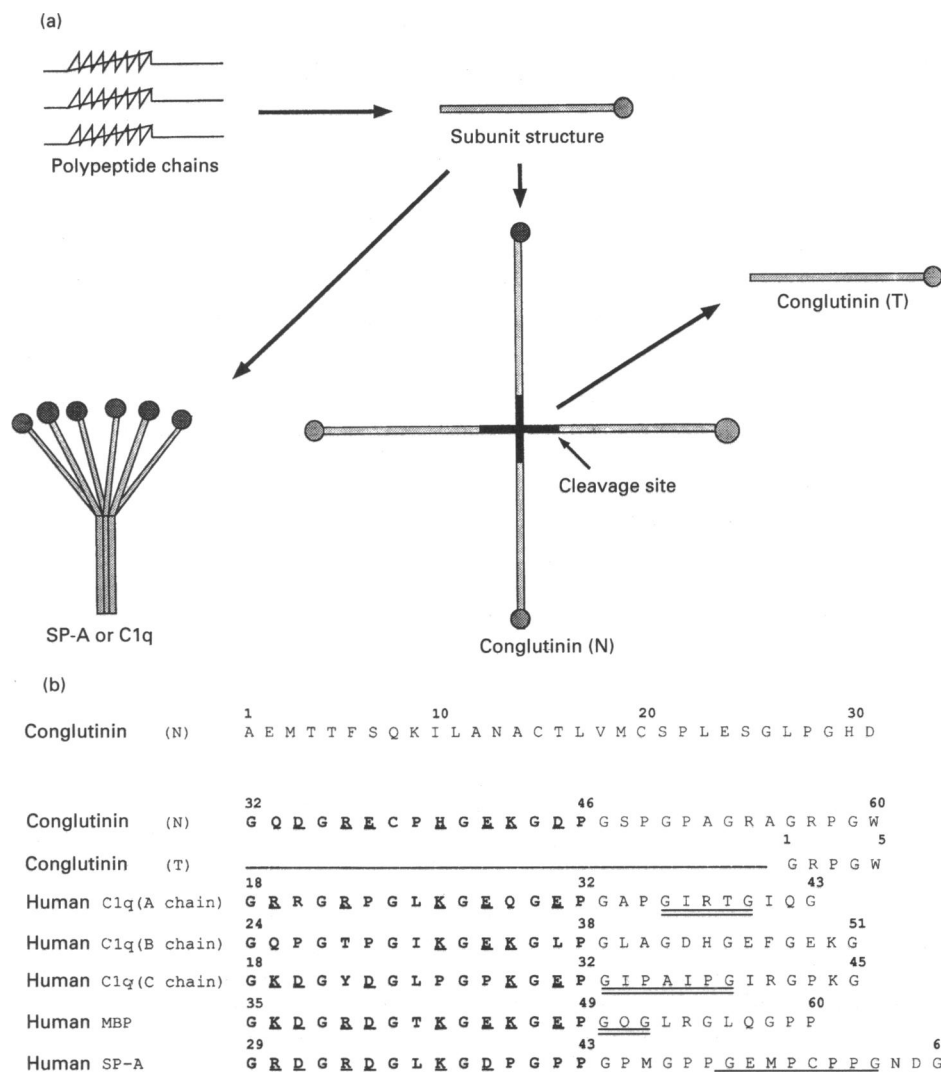| | | 29 | | | | | | | | | | | | | 43 | | | | | | | | | | | | | 60 |
|---|---|---|
| Human SP-A | G R D G R D G L K G D P G P P G P M G P P G E M P C P P G N D G |

Figure 1   Schematic illustration of the assembly of SP-A, C1q and conglutinin (N) subunits from three polypeptide chains (a) and the N-terminal amino acid sequences (b) of the two forms of conglutinin, conglutinin (N) and conglutinin (T)

(a) Polypeptide chain contains a collagen-like region (represented by zigzag lines) and non-collagenous sequence (represented by straight horizontal lines). The ultrastructures of SP-A, C1q and conglutinin (N) are based on the appearance of these molecules in electron microscopy [22]. The likely mechanism by which conglutinin (T) is generated from the normal form and the cleavage site on conglutinin (N) are also shown. (b) The figure also shows the amino acid sequence of human C1q (A chain, B chain, C chain), human MBP and human SP-A from a region close to the bend in the middle of the collagen region. Charged amino acids are underlined and the proposed binding site is shown in bold. The regions in C1q, MBP and SP-A in which there is an interruption to the Gly-Xaa-Yaa triplet are double underlined.

[3,6] and the purified receptor, as analysed by SDS/PAGE, is shown in Figure 2 (lane 1). Human SP-A was purified as described previously [1]. Bovine conglutinin was purified (S. B. Laursen, S. Thiel and J.-C. Jensenius, unpublished work). In brief, conglutinin was precipitated from bovine serum with 3.5 % (w/v) poly(ethylene glycol) 6000 (PEG 6000) (Sigma). The pellet was washed twice with 5 mM sodium barbitone buffer containing 145 mM NaCl, 0.05 % Tween 20, 2 mM CaCl$_2$ (buffer A) and 3.5 % (w/v) PEG 6000. The pellet was finally resuspended in buffer A and was loaded on to an N-acetyl-D-glucosamine–Sephacryl S-300 column, pre-equilibrated with buffer A. N-Acetyl-D-glucosamine was coupled to Sephacryl S-300 beads by the method of Fornstedt and Porath [13]. The column was washed extensively with buffer A and conglutinin-containing fractions were eluted with buffer A containing 2 mM N-acetyl-D-glucosamine. The conglutinin-containing fractions were dialysed against buffer A and reloaded on to the N-acetyl-D-glucosamine–Sephacryl S-300 column. Conglutinin-containing fractions were then eluted with buffer A containing 10 mM EDTA instead of 2 mM CaCl$_2$. These fractions were absorbed on a rabbit anti-(bovine immunoglobulin)–Sepharose column and then loaded on to a Sephacryl S-400 gel-filtration column in buffer A. Two different forms of conglutinin were eluted from the gel-filtration column.

Radio-iodination of collectin receptor

Radio-iodination was carried out by the Iodogen method of Fraker and Speck [14]. The specific radioactivity of the labelled collectin receptor was 4×10^7 c.p.m./μg of protein, and the autoradiograph of the radiolabelled receptor, as analysed on SDS/PAGE, is shown in Figure 2 (lane 2).
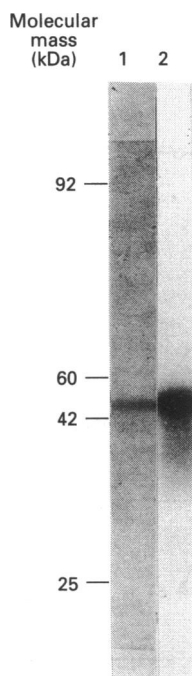
Figure 2   SDS/PAGE analysis of unlabelled purified collectin receptor (lane 1) and autoradiograph of the dried SDS/PAGE gel of radiolabelled collectin receptor (lane 2)

Solid-phase binding assays

Microtitre plate wells were coated with 100 µl of different forms of the purified conglutinin (50 µg/ml) or BSA (50 µg/ml) in PBS. Non-specific binding sites were blocked with 300 µl of BSA (5 mg/ml) in PBS. Radio-iodinated collectin receptor (100 µl; 2.3×10^5 c.p.m.) in 10 mM sodium phosphate buffer (pH 7.4) containing 0.1 % Emulphogene was loaded on to the protein-coated wells and incubated for 1 h at room temperature. After extensive washing with loading buffer the bound radioactivity was eluted with 200 µl of 4 M NaOH and the eluted radioactivity was measured.

In further sets of experiments microtitre plate wells were coated with C1q (10 µg/ml) in PBS. Radio-iodinated collectin receptor (100 µl; 2.3×10^5 c.p.m.) was incubated for 1 h at room temperature in the presence of serial dilutions of C1q collagen stalks (100 µl; maximum amount 13 pmol), SP-A (100 µl; maximum amount 4 pmol), the truncated form of conglutinin [conglutinin (T)] (100 µl; maximum amount 63 pmol), the normal form of conglutinin [conglutinin (N)] (100 µl; maximum amount 14 pmol), C1q (100 µl; maximum amount 16 pmol) or BSA (100 µl; maximum amount 63 pmol) in 10 mM sodium phosphate buffer containing 0.1 % (w/v) Emulphogene (pH 7.4). The mixture was loaded on to the C1q-coated microtitre plate wells and incubated for 1 h at ambient temperature. After extensive washing with loading buffer the bound radioactivity was eluted with 200 µl of 4 M NaOH and the eluted radioactivity was measured.

N-terminal amino acid sequence analysis

Samples for sequence analysis were run on SDS/PAGE under reducing conditions and electroblotted on to Problott membrane (Applied Biosystems, Warrington, U.K.) in a Bio-Rad mini

Trans-Blot electrophoretic transfer cell. The blots were stained with Coomassie Brilliant Blue and the bands corresponding to different forms of conglutinin were excised and sequenced using an Applied Biosystems 470A protein sequencer and Applied Biosystems 120A analyser.

RESULTS

In Figure 3(a) is shown SDS/PAGE analysis, under reducing conditions, of the purified conglutinin-containing fractions at different stages of purification. Two forms of the conglutinin polypeptide chain, molecular masses 40 kDa and 44 kDa, were co-purified from the N-acetyl-D-glucosamine–Sephacryl S-300 column (Figure 3a; lane 4). These two species of conglutinin were separated by Sephacryl S-400 gel filtration (Figure 3b) and the appearance of the two species of conglutinin polypeptide, after the gel-filtration step, is shown in Figure 3(a) (lanes 2 and 3). The elution profiles of the two forms of conglutinin, in comparison with standard proteins, by Sephacryl S-400 gel filtration shows that the form containing the 44 kDa polypeptide chains, conglutinin (N), elutes between the void volume and the
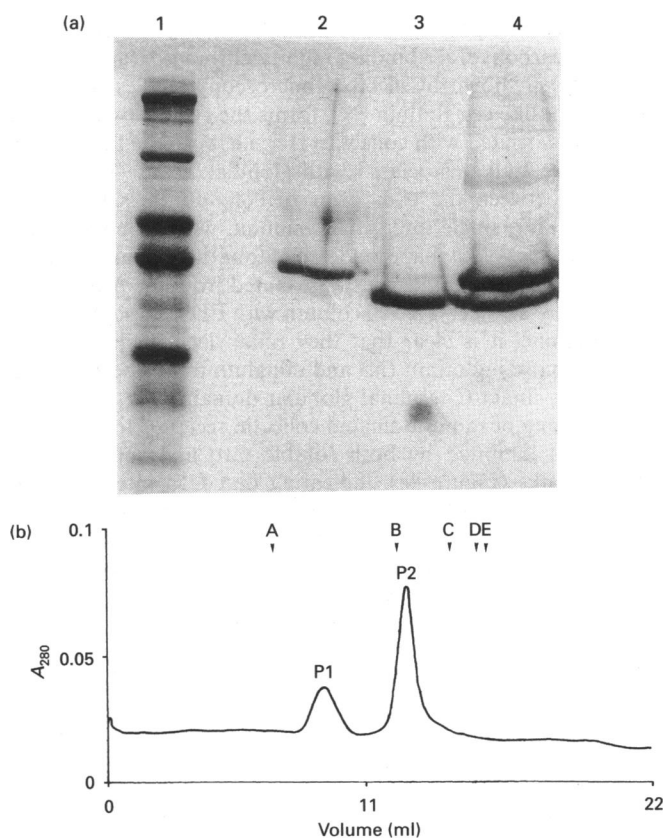

Figure 3   SDS/PAGE analysis of different forms of conglutinin (a) and elution positions of purified conglutinin (N) (P1) and conglutinin (T) (P2) from Sephacryl S-400 gel filtration (b)

(a) Lane 1, molecular-mass standards: from top to bottom, myosin (212 kDa), phosphorylase b (92 kDa), catalase (60 kDa), ovalbumin (42 kDa) and carboxypeptidase (25 kDa); lane 2, purified conglutinin (N) under reducing conditions (material from peak P1, Figure 2b); lane 3, purified conglutinin (T) under reducing conditions (material from peak P2, Figure 2b); and lane 4, fractions containing both forms of conglutinin after the N-acetyl-D-glucosamine–Sephacryl S-300 column purification step, under reducing conditions. (b) Elution positions of molecular-mass markers: A, Blue Dextran (2000 kDa); B, thyroglobulin (669 kDa); C, ferritin (440 kDa); D, catalase (232 kDa); E, aldolase (158 kDa) are indicated by arrows.

670 kDa thyroglobulin marker. This is consistent with the very elongated shape of conglutinin. The form containing 40 kDa polypeptides, conglutinin (T), elutes at a slightly smaller size than thyroglobulin, indicating that it also is multimeric and elongated.

The N-terminal amino acid sequence analysis of these two forms of conglutinin revealed that the lower-molecular-mass species conglutinin (T) is a truncated form of the higher-molecular-mass species conglutinin (N) (Figure 1b). The origin of conglutinin (T) in serum is not clear, but it is likely to arise by proteolytic cleavage of conglutinin (N) (Figure 1a). Cleavage is at an Ala-Gly bond. Leucocyte elastase may be a candidate protease for cleavage, but there is no firm evidence of this as yet. The N-terminal sequence of conglutinin (T) (Figure 1b) indicates that 55 residues are missing from the N-terminus of each polypeptide chain compared with conglutinin (N). This missing portion, if it remained intact, could be in the form of four segments of triple helices, i.e. $4 \times 3 \times 55$ amino acid residues, or about 65–70 kDa. It may also, however, be further proteolysed to small fragments, or its polypeptides may dissociate. A thorough search of the fractions eluted from gel filtration (Figure 3b) failed to reveal any material corresponding to this other piece of conglutinin, or to sub-fragments of it. It is clear, therefore, that it dissociates from conglutinin (T) at an earlier stage (i.e. immediately after proteolysis, or during PEG 6000 precipitation, or during carbohydrate-binding affinity chromatography).

It has been shown by electron-microscopy studies that conglutinin (T), like conglutinin (N), forms the typical two-domain structure associated with collectins [15], i.e. both the forms have collagen-like stalks associated with globular heads. The major difference between the two forms of conglutinin is that conglutinin (T) is a single three-chain subunit, whereas conglutinin (N) forms a cross-like shape made up of four three-chain subunits (Figure 1a). As both molecules were eluted from the N-acetyl-D-glucosamine–Sephacryl S-300 column with EDTA and N-acetyl-D-glucosamine, it is clear that they retain lectin activity. This indicates that conglutinin (N) and conglutinin (T) both have a functionally intact C-terminal globular domain.

The binding of radio-iodinated collectin receptor to immobilized C1q is inhibited by both soluble C1q and soluble C1q-collagen stalks (Figure 4a), indicating that C1q interacts with collectin receptor through the collagen domain. This is consistent with the earlier reports that the binding of C1q and SP-A to cells expressing their receptor takes place through the collagen domain [16–19]. Consistent with earlier data [2], SP-A also inhibits C1q–collectin-receptor interaction (Figure 4a).

It has been shown previously that binding of SP-A and C1q to collectin receptor is dependent on ionic strength [1,3]. On the basis of binding characteristics and sequence comparison between the four collectins, i.e. MBP, SP-A, C1q and conglutinin, we had previously suggested that the region close to the collagen triple-helix bend region, a region of charged residues (Figure 1b), may be the binding site for SP-A or C1q with the collectin receptor [1]. Conglutinin (N) contains a similar region of sequence (Figure 1b), but this is absent from conglutinin (T). The use of these two forms of conglutinin, therefore, may provide evidence to support the suggested receptor-binding site.

Collectin receptor does bind to conglutinin (N) but not to conglutinin (T) (Table 1). Similarly the binding of collectin receptor to solid-phase immobilized C1q was inhibited by soluble C1q, and soluble conglutinin (N), but conglutinin (T) had only a minor effect on the interaction between collectin receptor and C1q (Figure 4b). The slight inhibition of binding of collectin receptor to C1q seen with higher concentrations of conglutinin (T) is probably due to minor contamination with conglutinin (N)
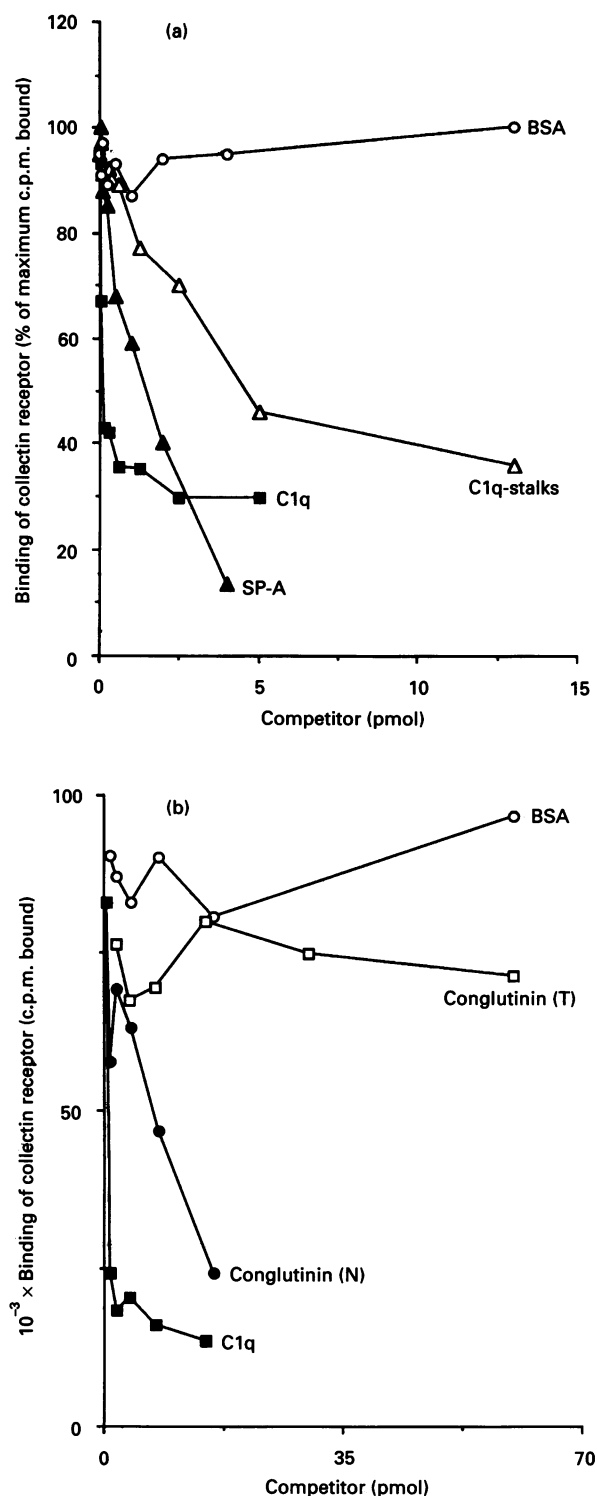

Figure 4   Binding characteristics of the collectin receptor

(a) Binding of radio-iodinated collectin receptor to solid-phase immobilized C1q in the presence of serial dilutions of C1q (■), C1q-collagen stalks (△), SP-A (▲) or BSA (○). Details are given in the Materials and methods section. Results are of three experiments, with the average of triplicate experimental points shown. (b) Binding of radio-iodinated collectin receptor to solid-phase immobilized C1q in the presence of serial-dilutions of C1q (■), conglutinin (N) (●), conglutinin (T) (□) and BSA (○). Details are given in the Materials and methods section. Results are of three experiments, with the average of triplicate experimental points shown. To calculate the amount of collectins in terms of moles the molecular masses of collectins were assumed to be: C1q, 460 kDa; C1q-collagen stalks, 190 kDa; SP-A, 600 kDa; conglutinin (N), 528 kDa; conglutinin (T), 120 kDa.

Table 1   Binding of radiolabelled collectin receptor to solid-phase immobilized conglutinin (N), conglutinin (T) and BSA

Details are given in the Materials and methods section. Results of two experiments each with the average of six experimental points and standard deviation are shown.

Ligand	Binding (c.p.m. bound ± S.D.)
Conglutinin (N)	9700 ± 713
Conglutinin (T)	979 ± 284
BSA	562 ± 121

(Figure 3a; lane 3). Conglutinin (T) differs from conglutinin (N) in that each polypeptide lacks the first 55 amino acids of the normal sequence (Figure 1b). The absence of this sequence clearly prevents interaction with receptor, as shown here, and also prevents formation of the four-subunit structure found in normal conglutinin ([15]; Figure 1a).

DISCUSSION

We have previously shown that binding of SP-A and C1q to U937 cells and purified collectin receptor is dependent on ionic strength, implying involvement of charged amino acids in the interaction of collectins with collectin receptor [1,3]. The binding characteristics of interaction of SP-A with an SP-A receptor on human cell lines have not been extensively investigated, but a number of workers have characterized the binding of SP-A to rat alveolar type-II cells and rat alveolar macrophages. Rice et al. [18] and Wright et al. [20] provided evidence that the interaction of SP-A with receptor did not involve the lectin domain. Kuroki et al. [19] also showed that mannose or α-methylmannoside or concanavalin A did not effect the binding of radio-iodinated rat SP-A to alveolar type-II cells, nor did these substances affect the SP-A-induced inhibition of surfactant secretion. They also showed that chemical modification of basic amino acids in SP-A resulted in inhibition of binding of radio-iodinated SP-A to alveolar type-II cells and in inhibition of the biological activity of SP-A. This observation is consistent with the binding site suggested by us (Figure 1b). Interaction of dog SP-A with dog alveolar macrophages and polymorphs is destroyed by collagenase treatment [21]. In these earlier studies, the receptor to which SP-A was binding was uncharacterized, but it is now clear that SP-A binds to U937 cells and A549 cells via the collectin receptor [1,8]. The results presented in this paper show that the binding of collectin receptor to collectins takes place through the collagen domain and the binding site for collectin receptor on collectins consists of a region with a high proportion of charged

residues, which is consistent with the characteristics of SP-A or C1q interaction with different cell types as discussed above.

Binding of conglutinin (or other collectins) to the collectin receptor may be mediated by a linear sequence motif in the conglutinin polypeptide chain, or may require features of distribution of charge around a single collagen triple helix, or even around several helices, since in all the collectins the proposed binding site is in a region where several helices are in proximity. Whatever the requirement, the receptor–ligand interaction is clearly mediated via the N-terminal region of conglutinin, consistent with the earlier proposal for the binding site [1]. These results also indicate that collectin receptor is not a general receptor for collagen-like structures, but requires a specific charge distribution.

REFERENCES

1   Malhotra, R., Haurum, J., Thiel, S. and Sim, R. B. (1992) Eur. J. Immunol. **22**, 1437–1445
2   Malhotra, R., Thiel, S., Reid, K. B. M. and Sim, R. B. (1990) J. Exp. Med. **172**, 955–959
3   Malhotra, R. and Sim, R. B. (1989) Biochem. J. **218**, 625–631
4   McCauliffe, D. P., Lux, F. A., Lieu, T.-S., Sanz, I., Hanke, J., Newkirk, M. M., Bachinski, L. L., Itoh, Y., Siciliano, M. J., Reichlin, M., Sontheimer, R. D. and Capra, J. D. (1990) J. Clin. Invest. **85**, 1379–1391
5   Rokeach, L. A., Haselby, J. A., Meilof, J. F., Smeenk, R. J. T., Unnasch, T. R., Greene, B. M. and Hoch, S. O. (1991) J. Immunol. **147**, 3031–3039
6   Malhotra, R., Willis, A. C., Jensenius, J. C., Jackson, J. and Sim, R. B. (1993) Immunology **78**, 341–349
7   Peerschke, E. I. B., Malhotra, R., Ghebrehiwet, B., Reid, K. B. M., Willis, A. C. and Sim, R. B. (1993) J. Leukocyte Biol. **53**, 179–184
8   Malhotra, R., Lu, J., Thiel, S., Jensenius, J.-C., Willis, A. C. and Sim, R. B. (1992) Immunobiology **184**, 437–438
9   Pison, J., Wright, J. E. and Hawgood, S. (1992) Am. J. Physiol. **262**, L412–416
10   Laemmli, U. K. (1970) Nature (London) **227**, 680–685
11   Fairbanks, G., Steck, T. L. and Wallach, D. F. H. (1971) Biochemistry **10**, 2602–2617
12   Reid, K. B. M. (1981) Methods Enzymol. **80**, 16–25
13   Fornstedt, N. and Porath, J. (1975) FEBS Lett. **57**, 187–190
14   Fraker, P. J. and Speck, J. C., Jr. (1978) Biochem. Biophys. Res. Commun. **80**, 849–857
15   Lu, J., Wiedemann, H., Holmskov, U., Thiel, S., Timpl, R. and Reid, K. B. M. (1993) Behring Inst. Mitt., in the press
16   Sobel, A. T. and Bokisch, V. A. (1975) Fed. Proc. Fed. Am. Soc. Exp. Biol. **34**, 965
17   Arvieux, J., Reboul, A., Bensa, J. C. and Colomb, M. G. (1984) Biochem. J. **218**, 547–555
18   Rice, W. R., Ross, G. F., Singleton, F. W., Dingle, S. and Whitsett, J. A. (1987) J. Appl. Physiol. **63**, 692–698
19   Kuroki, Y., Mason, R. J. and Voelker, D. R. (1988) J. Biol. Chem. **263**, 17596–17602
20   Wright, J. R., Wager, R. E. and Hawgood, S. (1987) J. Biol. Chem. **262**, 2888–2894
21   Weber, H., Heilmann, P., Meyer, B. and Maier, K. L. (1990) FEBS Lett. **270**, 90–94
22   Thiel, S. and Reid, K. B. M. (1989) FEBS Lett. **250**, 78–84